Amino acid sequence of an extracellular, phosphate-starvation-induced ribonuclease from cultured tomato (Lycopersicon esculentum) cells

Jost, Wolfgang; Bak, Henk; Glund, Konrad; Terpstra, Peter; Beintema, Jacob

Published in:
European Journal of Biochemistry

DOI:
10.1111/j.1432-1033.1991.tb15978.x

Eur. J. Biochem. *198*, 1–6 (1991)

0014295691003O1Y

Amino acid sequence of an extracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*) cells

Wolfgang JOST[1], Henk BAK[3], Konrad GLUND[1], Peter TERPSTRA[2] and Jaap J. BEINTEMA[2]

[1] Martin-Luther-Universität Halle-Wittenberg, Institut für Biochemie, Pflanzenbiochemie, Halle, Federal Republic of Germany
[2] Biochemisch Laboratorium, Rijksuniversiteit and [3] Eurosequence B. V., Groningen, The Netherlands

(Received November 5, 1990) – EJB 90 1298

The primary structure of an extracellular ribonuclease (RNase LE) from P_i-depleted media of cultured cells of *Lycopersicon esculentum* L. cv. *Lukullus* has been determined. This was carried out by analysis of peptides isolated after enzymatic and chemical cleavage of the reduced and *S*-ethylpyridylated protein.

RNase LE consists of 205 amino acid residues and has a molecular mass of 22666 Da and an isoelectric point of 4.24. The enzyme contains 10 half-cystines. There are no potential *N*-glycosylation sites in the sequence.

The sequence of RNase LE is homologous with those of self-incompatibility proteins of several higher plant species and with those of a number of fungal RNases. The sequence similarity with the family of self-incompatility proteins is greater than with the fungal RNases, suggesting that the self-incompatibility proteins arose from ancestral RNase by gene duplication after the divergence of higher plants and fungi. Two pentapeptide sequences, i.e. HGLWP and KHGTC (or KHGSC), are present at identical positions in all the aligned proteins, suggesting that they contribute to the active site.

RNA metabolism, which includes processing, turnover and degradation of cellular RNA, has emerged as a complex and important determinant of gene expression in living cells. Much of our current knowledge on enzymes hydrolyzing RNA (RNases) and their *in vivo* functions comes from studies with *Escherichia coli* [1]. With respect to structural characterization and regulation of RNases from plant origin, knowledge lags far behind that already well established for the enzymes from fungi [2] and mammals [3].

A number of physiological functions in plants were found to be associated with changes in nucleolytic activities (see [4] and [5] for reviews). More recent studies on nucleolytic and ribonucleolytic enzymes have shown that their expression is regulated by hormones [6], light and senescence [7], plant development [8], aging and wounding [9] or nutrient starvation [10, 11] (see below). Investigation of these enzymes may lead to the identification of regulatory elements and signals for plant gene expression in that field.

Correspondence to W. Jost, Martin-Luther-Universität, Institut für Biochemie, Pflanzenbiochemie, Weinbergweg 16a, O-4050 Halle, Federal Republic of Germany

Abbreviations. RNase, ribonuclease; RNase LE, extracellular *Lycopersicon* RNase (extracellular RNase from cultured tomato cells); EtPd-RNase LE, reduced and *S*-ethylpyridylated RNase LE; DABITC, 4-(*N*,*N*-dimethylamino)azobenzene 4′-isothiocyanate; PhSCN, phenylisothiocyanate.

Enzymes. RNase (EC 3.1.27.1); α-chymotrypsin (EC 3.4.21.1); trypsin (EC 3.4.4.4); thermolysin (EC 3.4.24.4); endoproteinase Glu-C from *Staphylococcus aureus* V8 (EC 3.4.21.19); carboxypeptidase A (EC 3.4.17.1); carboxypeptidase Y (EC 3.4.16.1).

Note. The novel amino acid sequence data published here have been submitted to the EMBL data library.

Among the better studied fungal RNases which cleave RNA via nucleotide 2′,3′-(cyclic)phosphates as intermediates, the primary structures of RNase T2 from *Aspergillus oryzea* [12], RNase M from *Aspergillus satoi* [13] and RNase Rh from *Rhizopus niveus* [14] have been determined. There is a significant similarity between these extracellular RNases and glycoproteins of tobacco and other plants, which are encoded by the S gene, confer self-incompatibility [15] and were identified as ribonucleases [16].

Previously, we have identified from cultured tomato (*Lycopersicon esculentum*) cells, an extracellular 22-kDa RNase which was shown to cleave 3′ – 5′ internucleotide linkages of RNA via nucleotide 2′,3′-(cyclic)phosphates as intermediates, and to release preferentially 3′-purine mononucleotides from its substrates. The enzyme was found to be synthesized and secreted from the cells if they were subjected to conditions of phosphate depletion [11, 17].

As a basis for further studies on the chemistry and molecular biology of these enzymes, we present here the complete primary structure of the extracellular tomato RNase.

MATERIALS AND METHODS

Materials

Tosylphenylalanylchloromethane-treated-trypsin and α-chymotrypsin were purchased from Worthington Biochem. Corp., endoproteinase Glu-C from *Staphylococcus aureus* V8, thermolysin, carboxypeptidase A and carboxypeptidase Y were obtained from Boehringer. Reagents used for automated sequencing were from Applied Biosystems. All other chemicals were of the highest grade commercially available.

Preparation of RNase LE

RNase LE was purified from phosphate-depleted culture media of *L. esculentum* L. cv. *Lukullus* cells as described [11]. Preparations showing a single spot on two-dimensional electrophoresis [18], and eluting as a single peak on reversed-phase HPLC, were used.

Amino acid analysis

Protein or peptides were hydrolyzed in 6 M twice-distilled HCl *in vacuo* at 110°C for 24 h. The hydrolysates were analyzed with a Hewlett Packard AminoQuant (1090 M) amino acid analyzer.

Reduction and S-ethylpyridylation

RNase LE was reduced and *S*-ethylpyridylated by treating 4 mg protein dissolved in 1 ml 50% pyridine with 5 µl tributylphosphine under nitrogen for 30 min at room temperature, followed by treatment with 5 µl 4-vinylpyridine for 30 min under the same conditions. The derivative (EtPd-RNase LE) was separated from the other reagents by reversed-phase HPLC.

Enzymatic digestions

EtPd-RNase LE or peptides derived from it were dissolved in 8 M urea, diluted four times with 0.1 M ammonium bicarbonate, pH 8.0, and digested in several batches by trypsin, endoproteinase Glu-C, thermolysin or α-chymotrypsin for 4–6 h at 37°C at a substrate/enzyme ration of 100:1 (by mass). EtPd-RNase LE was digested with carboxypeptidase A (prepared as described by Ambler [19]) or carboxypeptidase Y in 0.1 M ammonium bicarbonate, pH 8.5 for 5 h at 37°C at a substrate/enzyme ratio of 3:1 (by mass). Free amino acids were identified by dansylation [20] or with a Hewlett Packard AminoQuant (1090 M) amino acid analyzer.

Cyanogen bromide cleavage

0.3 mg EtPd-RNase LE in 30 µl 70% formic acid was treated with 0.6 mg CNBr in the dark, under nitrogen, and at room temperature for 24 h [21]. The reaction mixture was evaporated to dryness before further use.

Selective acid hydrolysis of Asp-Pro bonds

Partial acid cleavage of EtPd-RNase LE was performed in 75% formic acid for 48 h at 37°C [22].

HPLC separation of protein and peptides

Reversed-phase HPLC was carried out with a Nucleosil 10 C_{18} column (4.6 mm × 300 mm) at a flow rate of 1 ml/min. The gradients were prepared using a Waters model 680 gradient controller with two 6000A pumps. System 1: a linear gradient of 0–60% acetonitrile in 0.1% trifluoroacetic acid, run at 1.1% acetonitrile/min. System 2: like system I, but at 0.33% acetonitrile/min. System 3: a linear gradient of 0–60% acetonitrile in 0.1% ammonium acetate, pH 6.0, run at 1.1% acetonitrile/min. System 1 was used for chromatography of RNase LE and EtPd-RNase LE. Peptides were usually separated in the same system, and those not resolved were rechromatographed in system 2 or system 3.

Sequence analysis

Automated Edman degradation was performed on an Applied Biosystems 477A pulse-liquid sequencer with an on-line 120A phenylthiohydantoin analyzer of the sequence facility of the Institute Bioson and Eurosequence BV. Special programs supplied with the instrument were used at several proline residues in the sequence with an *o*-phthalic acid derivatization cycle if there is an N-terminal proline residue and a double-cleavage cycle to cleave the peptide bond on the C-terminal side of this residue.

Manual sequencing with the 4-(*N,N*-dimethylamino)azobenzene/phenylisothiocyanate (DABITC/PhSCN) double-coupling procedure of Chang et al. [23] was used for identifying peptides and checking them for purity.

Similarity in amino acid sequence between RNase LE and other proteins was tested by scanning the SWISS-PROT database (release 12; November 1989) with the program FASTa [24].

Peptide nomenclature

Peptides are numbered according to their final alignment in the sequence. The prefixes T, G, TH, CH, A, and C indicate the tryptic, endoproteinase Glu-C, thermolytic, chymotryptic, acid cleavage and CNBr peptides, respectively.

RESULTS

Sequencing strategy

EtPd-RNase LE was cleaved by trypsin, chymotrypsin, thermolysin, CNBr, endoproteinase Glu-C and acid treatment. Fig. 1 summarizes the cleavage procedure used and the peptides obtained and sequenced. For ease of survey, only fragments indispensable for the sequence determination are shown.

Determination of the N-terminal sequence

The N-terminal sequence of the protein was deduced from a sequence run of the unmodified protein (28 residues). At three positions, no amino acid residue could be identified. Later, half-cystines could be assigned to these positions from a sequence run on an equimolar mixture of EtPd-RNase LE and the CNBr-cleaved peptide C, and from the sequence run on peptide T-2, respectively. In addition, the N-terminus was confirmed by the N-terminal sequence and amino acid analysis of T-1.

Determination of the C-terminal sequence

As indicated in Table 1, T-8 is the only tryptic peptide which contains neither Lys nor Arg. This, together with the fact that during automated Edman degradation of this peptide, all phenylthiohydantoins abruptly reached background levels after step 10, clearly reveals Phe205 in T-8 as the C-terminal residue.

This conclusion has been confirmed by carboxypeptidase digestion of EtPd-RNase LE. Carboxypeptidase A liberated only Phe (0.8 mol/mol protein after 5 h digestion at 37°C). After digestion with carboxypeptidase Y, Phe, Thr and Pro were found, in addition to C-terminal amino acids from the

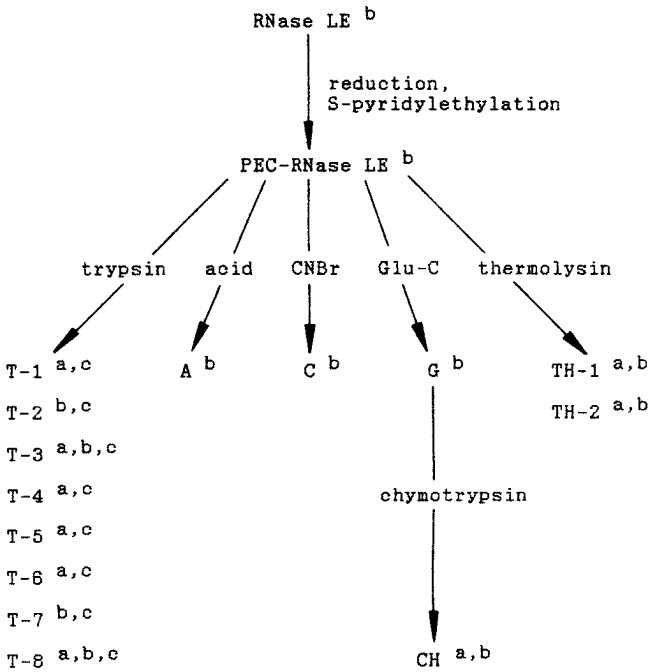

Fig. 1. *Sequencing strategy.* RNase LE was modified and cleaved as indicated. Peptides were chromatographed in the following HPLC systems: T, system 1; A, systems 1 and 2; C, systems 1 and 2; G, systems 1 and 2; TH, systems 1 and 3; CH, system 1 (a) sequenced manually for at least three steps; (b) automated Edman degradation; (c) amino acid analysis. PEC-RNase LE, EtPd-RNase LE

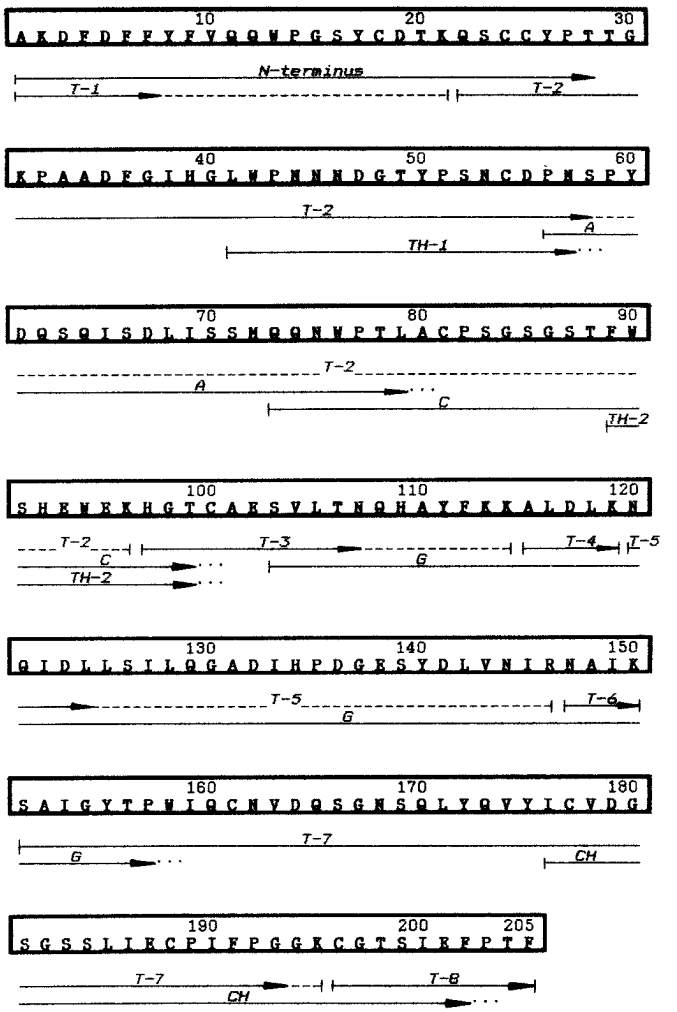

Fig. 2. *Amino acid sequence of RNase LE.* Prefixes T, G, TH, CH, A, and C indicate tryptic, endoproteinase Glu-C, thermolytic, chymotryptic, acid cleavage and CNBr peptides, respectively. (→) Sequenced (– – – –) covered by amino acid analysis; (· · · ·) length of fragment unknown

enzyme itself, that also appeared in a blank digestion experiment without EtPd-RNase LE.

Sequencing of EtPd-RNase LE fragments

Most information about the primary structure of RNase LE (Fig. 2) was obtained from sequencing peptides T-7 and G which could be run to degradation steps 43 and 55, respectively. This, together with the N- and C-terminal peptides (see above), covered about 63% of the total protein sequence.

Additional peptides from other fragmentations (Fig. 1) were isolated, screened for appropriate N-termini by the DABITC/PhSCN method and, subsequently, run in the sequencer to obtain overlaps.

As shown in Fig. 2, T-2 overlapped the N-terminus as well as fragment A. The identity of residues 41 – 59 in T-2 was ambiguous due to a low signal-to-noise ratio. This sequence was confirmed by the sequence of TH-1, which also provided the overlap with peptide A.

Peptides A and C could not be purified to homogeneity. In order to get unambiguous results, the N-termini of contaminating sequences were blocked by the introduction of *O*-phthalic acid and double-cleavage cycles for peptide A at step 1 and for peptide C at steps 5 and 10. In this way, the sequence of peptide A was determined until an overlap with peptide C was reached. Despite a relatively high initial yield (Fig. S1), the sequence of peptide C became uncertain after step 19. However, the sequence of TH-2 provided unambiguous proof of this part of the sequence and the overlap with peptide T-3. The N-terminal sequence of peptide C was also provided by the sequencer run on the equimolar mixture of this peptide with intact EtPd-RNase LE mentioned above.

Finally, unidentified residues between peptides TH-2 and G, as well as between T-7 and T-8, were covered by T-3 and CH, respectively.

Most parts of RNase LE were sequenced at least twice using different fragments, leading to unambiguous results. Additional evidence was provided by the results of manual sequencing and amino acid analyses of RNase LE and tryptic peptides (Table 1).

Primary structure of RNase LE

In Fig. 2, the established amino acid sequence is presented. RNase LE consists of 205 amino acid residues. The amino acid composition derived from the sequence is in agreement with results of amino acid analysis of the whole protein (not shown). Values of 22666 Da for molecular mass and of 4.24 for isoelectric point, calculated from the sequence with PC-GENE program, fit well with those determined by SDS/PAGE and IEF [11]. The enzyme contains 10 half-cystines. There are no potential *N*-glycosylation sites in the sequence.

Table 1. *Amino acid composition of tryptic peptides from EtPd-RNase LE*
Values in parentheses are from the established amino acid sequence; half-cystine and tryptophan were not determined; n.d., not determined

Amino acid	Amino acid content of							
	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8
	mol/mol							
Asx	3.6 (3)	11.7 (11)	0.9 (1)	1.0 (1)	5.8 (6)	1.1 (1)	4.4 (4)	
Glx	2.0 (2)	8.4 (7)	2.0 (2)		3.0 (3)		5.5 (5)	0.9 (1)
Ser	2.0 (1)	11.0 (11)	1.2 (1)		1.7 (2)		6.3 (6)	0.9 (1)
His	0.1 (0)	n.d. (2)	1.6 (2)		0.3 (1)			
Gly	1.6 (1)	6.9 (6)	1.9 (1)		2.0 (2)		5.9 (6)	1.0 (1)
Thr	1.3 (1)	5.0 (5)	1.9 (2)				1.2 (1)	2.0 (2)
Ala	1.1 (1)	3.4 (3)	2.0 (2)	1.0 (1)	1.1 (1)	1.0 (1)	1.3 (1)	
Arg					0.4 (1)			
Tyr	1.7 (2)	3.4 (3)	0.5 (1)		0.5 (1)		2.5 (3)	
Val	0.6 (1)	0.1 (0)	0.8 (1)		0.9 (1)		2.7 (3)	
Met	0.1 (0)	1.3 (1)						
Ile	0.6 (0)	3.3 (3)			3.6 (4)	0.7 (1)	4.1 (5)	1.1 (1)
Phe	2.5 (4)	2.5 (2)	0.6 (1)				1.0 (1)	2.2 (2)
Leu	0.4 (0)	3.3 (3)	0.9 (1)	2.0 (2)	4.5 (4)		2.1 (2)	
Lys	1.8 (2)	2.5 (2)	1.5 (2)	0.7 (1)		0.8 (1)	1.0 (1)	
Pro	n.d. (1)	7.4 (8)	0.9 (0)		1.4 (1)		n.d. (3)	0.9 (1)

DISCUSSION

The present study was undertaken as a first step toward a chemical and molecular biological characterization of plant ribonucleases. The extracellular tomato ribonuclease under study is inducible upon phosphate starvation [11, 17], and this protein is most likely a member of a putative plant phosphate-starvation rescue system [25]. In order to obtain greater insight into the regulation of the biosynthesis of the protein, analysis of the genomic structure is required. The amino acid sequence presented here may facilitate the determination of this structure.

Our previous studies [11] have demonstrated that RNase LE splits RNA via nucleotide 2′,3′-(cyclic)phosphates preferentially adjacent to purine nucleotides in the 5′ position with a pH optimum of pH 5.5. The enzyme is insensitive to EDTA and has a p*I* of about 3.9. It is worth mentioning that we have characterized a cellular RNase from the same tomato cell cultures [26, 27] which is present in the vacuoles and has very similar properties including an identical N-terminus (the first seven amino acids of the vacuolar enzyme sequenced, not shown).

The amino acid sequence of RNase LE has significant similarity to those of the fungal RNases T2 from *A. oryzae* [12], RNase M from *A. satoi* [13] and RNase Rh from *R. niveus* [14], and those of several *S*-glycoprotein products of self-incompatibility genes from tobacco (*Nicotiana alata*) [15, 28] and petunia (*Petunia inflata*) [29]. The homologous nature of these fungal RNases and the *S*-glycoproteins has been recognized before and the latter proteins have been identified as ribonucleases [16]. The identity between the sequence of RNase LE and those of the fungal RNases is about 25%, and with those of the *S*-glycoproteins (Fig. 3), about 30%. RNase LE and the *S*-glycoproteins probably are orthologous gene products, which originate from an ancestral ribonuclease by gene duplication after the divergence of higher plants and fungi, as both proteins have been identified in the genus *Lycopersicon* [11, 30]. RNase LE is an acidic extracellular protein, while the *S*-glycoproteins are basic and pistil-specific.

The sequence of *S*-glycoproteins differ considerably between alleles of the same plant species. This complicates sequence comparison with other proteins. Kheyr-Pour et al. [28] compared the sequences of seven *S*-glycoproteins from tobacco and identified five conserved, a number of variable and five hypervariable sequence regions. The conserved regions are enclosed in boxes in the alignment with the RNase LE sequence in Fig. 3. It is evident that many residues in the conserved regions, and also a number of conserved residues in the variable regions, are also present in the sequence of RNase LE. The second and the third conserved boxes contain the pentapeptide sequences HGLWP and KHGTC or KHGSC. These are also conserved in the fungal RNase sequences [16], whereas the other three boxes are not.

Fungal RNases Rh [31, 32] and M [33, 34] both have two histidine residues and an acidic residue in their active sites. The two histidine residues have been identified by carboxymethylation [31, 33] and are those present in the two conserved pentapeptide sequences mentioned above. This suggests a catalytic mechanism similar to that proposed for bovine pancreatic ribonuclease [35].

Watanabe et al. [13] restricted the possible candidates for the acidic residue in the active site to Asp44 or Glu93 (Fig. 3; RNase LE sequence numbering). More recently, Glu93 has been identified as this active-site residue in RNase Rh by site-directed mutagenesis [36]. The replacement of Asp44 by Asn in RNase Rh does not abolish enzyme activity [36], but only decreases the preference for adenine residues in this purine-specific RNase (M. Irie, personal communication). This observation is in agreement with the guanine preference of RNase LE [11] and the presence of asparagine at position 44 in this enzyme (Fig. 3).

Glu93 is conserved in the fungal RNases, in RNase LE, and in most *S*-glycoproteins. However, *S*-glycoproteins SF11 and Sz have glutamine at this position [28]. The level of RNase activity in these two allele products has not yet been reported.

RNase LE contains 10 half-cystine residues. Eight of these are conserved residues in the *S*-glycoprotein sequences (Fig. 3) [28, 29]. Additional half-cystines occur in most *S*-glyco-

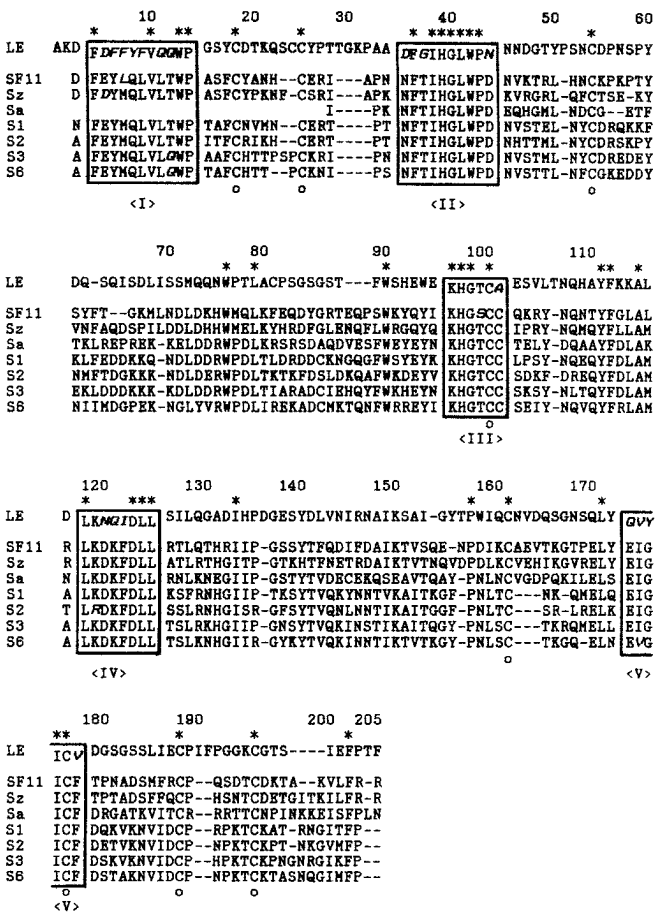

Fig. 3. *Alignment of the amino acid sequence of the extracellular RNase from* L. esculentum *(LE) with those of other RNases.* The program CLUSTAL [38] was used. RNase LE numbering is indicated. (*) Identical in all sequences aligned; (○) conserved half-cystines. SF11, Sz, Sa, S1 [28], S2, S3, S6 [39]: mature glycoproteins of *N. alata.* Boxes I – V are based on the conserved regions from [28]. Residues that are different from the conserved residues at those sites are given in italics

proteins at position 101 (except in one of the *Petunia* proteins) and in several *S*-glycoproteins at position 86. The RNase LE sequence has additional half-cystine residues at positions 24 (or 25) and 81, suggesting a disulfide bridge between these two positions in the sequence. Kawata et al. [12] have assigned two disulfide bridges in the sequence of RNase T2. The half-cystine residues participating in these two bridges may be homologous to the conserved residues at positions 54 and 100, and 161 and 196, in the alignment presented in Fig. 3. However, these very tentative assignments should be confirmed by chemical or X-ray-diffraction studies [37] of one or several members of this family of purine-preferring ribonucleases.

This work was supported by a fellowship to W. J. from the Federation of the European Biochemical Society. Additional support was obtained from grants of the Faculty of Mathematics and Sciences of the University of Groningen, the Groningen Biotechnology Center and the *Groninger Universiteitsfonds.* Authors are grateful to Helga Pechfelder for technical help, Piet Wietzes and Peter Jekel for performing amino acid analyses and the carboxypeptidase experiments, respectively, Thorsten Nürnberger for performing two-dimensional electrophoresis, and Steffen Abel for critical reading the manuscript.

REFERENCES

1. Deutscher, M. P. (1988) *Trends Biochem. Sci. 13*, 136 – 139.
2. Takahashi, K. & Moore, S. (1982) in *The enzymes* (Boyer, P. D., ed.) vol. 15, pp. 435 – 468, Academic Press, New York.
3. Beintema, J. J., Schüller, C., Irie, M. & Carsana, A. (1988) *Prog. Biophys. Mol. Biol. 51*, 165 – 192.
4. Wilson, C. M. (1982) in *Current topics in biological and medical research* (Ratazzi, M. C., Scandalios, J. G. & Whitt, G. S., eds) vol. 6, pp. 33 – 54, A. R. Liss, Inc., New York.
5. Farkas, G. L. (1982) in *Encyclopedia of plant physiology new series* (Parthier, B. & Boulter, D., eds) vol. 14B, pp. 224 – 262, Springer-Verlag, Berlin, Heidelberg, New York.
6. Brown, P. H. & Ho, T. D. (1987) *Eur. J. Biochem. 168*, 357 – 364.
7. Blank, A. & McKeon, T. A. (1989) *Proc. Natl Acad. Sci. USA 86*, 3169 – 3173.
8. Thelen, M. P. & Northcote, D. H. (1989) *Planta (Berl.) 179*, 181 – 195.
9. Isola, M. C. & Franzoni, L. (1986) *Plant Cell Physiol. 27*, 331 – 335.
10. Schroeder, J. (1977) *Arch. Biochem. Biophys. 182*, 488 – 496.
11. Nürnberger, T., Abel, S., Jost, W. & Glund, K. (1990) *Plant Physiol. 92*, 970 – 976.
12. Kawata, Y., Sakiyama, F. & Tamaoki, H. (1988) *Eur. J. Biochem. 176*, 683 – 697.
13. Watanabe, H., Naitoh, A., Suyama, Y., Inokuchi, N., Shimada, H., Koyama, T., Ohgi, K. & Irie, M. (1990) *J. Biochem. (Tokyo) 108*, 303 – 310.
14. Horiuchi, H., Yanai, K., Takagi, M., Yano, K., Wakabayashi, E., Sanda, A., Mine, S., Ohgi, K. & Irie, M. (1988) *J. Biochem. (Tokyo) 103*, 408 – 418.
15. Anderson, M. A., Cornish, E. C., Mau, S.-L., Williams, E. G., Hoggart, R., Atkinson, A., Bonig, I., Grego, B., Simpson, R., Roche, P. J., Haley, J. D., Penschow, J. D., Niall, H. D., Tregear, G. W., Coghlan, J. P., Crawford, R. J. & Clarke, A. E. (1986) *Nature 321*, 38 – 44.
16. McClure, B. A., Haring, V., Ebert, P. R., Anderson, M. A., Simpson, R. J., Sakiyama, F. & Clarke, A. E. (1989) *Nature 342*, 955 – 957.
17. Glund, K., Nürnberger, T., Abel, S., Jost, W., Preisser, J. & Komor, E. (1990) in *Proceedings of the VIIth international congress on plant tissue and cell culture*, Amsterdam (Nijkamp, H. J. J., Van der Plas, L. H. W. & Van Aartrijk, J., eds pp. 338 – 342, Kluwer Academic Publishers, Dordrecht, Boston, London.
18. O'Farrell, P. H. (1975) *J. Biol. Chem. 250*, 4007 – 4021.
19. Ambler, R. P. (1972) *Methods Enzymol. 25*, 143 – 154.
20. Hartley, B. S. (1970) *Biochem. J. 119*, 805 – 822.
21. Gross, E. (1967) *Methods Enzymol. 11*, 238 – 255.
22. London, M. (1977) *Methods Enzymol. 47*, 145 – 149.
23. Chang, J. Y. (1983) *Methods Enzymol. 91*, 455 – 466.
24. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl Acad. Sci. USA 85*, 2444 – 2448.
25. Goldstein, A. L., Baertlein, D. A. & Danon, A. (1989) *Plant Mol. Biol. Rep. 7*, 7 – 16.
26. Abel, S. & Glund, K. (1987) *Planta (Berl.) 172*, 71 – 78.
27. Abel, S., Krauss, G. J. & Glund, K. (1989) *Biochim. Biophys. Acta 998*, 145 – 150.
28. Kheyr-Pour, A., Bintrim, S. B., Ioerger, T. R., Remy, R., Hammond, S. A & Kao, T. (1990) *Sex. Plant Reprod. 3*, 88 – 97.
29. Ai, Y., Singh, A., Coleman, C. E., Ioerger, T. R., Kheyr-Pour, A. & Kao, T. (1990) *Sex. Plant Reprod. 3*, 130 – 138.
30. Mau, S.-L., Williams, E. G., Atkinson, A., Anderson, M. A., Cornish, E. C., Grego, B., Simpson, R. J., Kheyr-Pour, A. & Clarke, A. E. (1986) *Planta (Berl.) 169*, 184 – 191.
31. Sanda, A., Takizawa, Y. & Irie, M. (1985) *Chem. Pharm. Bull. (Tokyo) 33*, 4515 – 4521.
32. Watanabe, H., Sugiyama, K., Iwama, M., Futaki, R. & Irie, M. (1983) *J. Biochem. (Tokyo) 94*, 63 – 70.
33. Irie, M., Watanabe, H., Ohgi, K. & Harada, M. (1986) *J. Biochem. (Tokyo) 99*, 627 – 633.

34. Sanda, A., Takizawa, Y., Iwama, M. & Irie, M. (1985) *J. Biochem. (Tokyo) 98*, 125–132.
35. Blackborn, P. & Moore, S. (1982) in *The enzymes* (Boyer, P. D., ed.) vol. 15, pp. 317–433, Academic Press, New York.
36. Irie, M. (1990) in *Abstracts of the 2nd international meeting on structure, mechanism and function of ribonucleases*, Sant Feliu de Guixols, Spain, Abstract L21, p. 39.
37. Kurihara, H., Mitsui, Y., Nakamura, K. T., Wakabayashi, E., Ohgi, K. & Irie, M. (1989) *J. Mol. Biol. 206*, 791–792.
38. Higgins, D. G. & Sharp, P. M. (1988) *Gene 73*, 237–244.
39. Anderson, M. A., McFadden, G. I., Bernatzky, R., Atkinson, A., Orpin, T., Dedman, H., Tregear, G., Fernley, R. & Clarke, A. E. (1989) *Plant Cell 1*, 483–491

Supplementary material to

Amino acid sequence of an extracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*) cells

Wolfgang JOST, Henk BAK, Konrad GLUND, Peter TERPSTRA and Jaap J. BEINTEMA

N-terminus

```
Ala-Lys-Asp-Phe-Asp-Phe-Phe-Tyr-Phe-Val-Gln-Gln-Trp-Pro-Gly-Ser-
124  39  23  63  34  60  63  38  54  43  39  43  19  45  38  16

Tyr-Cys-Asp-Thr-Lys-Gln-Ser-Cys-Cys-Tyr-Pro-Thr
 14   +   8  14  13   9   5   +   +   4   6   1
```

T-1

```
Ala-Lys-Asp-Phe-Asp-Phe-Phe
 +   +   +   +   +   +   +
```

T-2

```
Gln-Ser-Cys-Cys-Tyr-Pro-Thr-Thr-Gly-Lys-Pro-Ala-Ala-Asp-Phe-Gly-
 64  74   +   +  48  52  42  49  28  10  37  41  49   6  27  30

Ile-His-Gly-Leu-Trp-Pro-Asn-Asn-Asn-Asp-Gly-Thr-Tyr-Pro-Ser-Asn-
 23   7  10  17   -  24  13  13   9   1   8   5   5   5   5   +

Cys-Asp-Pro-Asn-Ser
 -   +   +   -   +
```

T-3

```
His-Gly-Thr-Cys-Ala-Glu-Ser-Val-Leu-Thr-Asn
 +  601 691   +  882 378 373 486 471 455 303
```

T-4

```
Ala-Leu-Asp-Leu-Lys
 +   +   +   +   -
```

T-5

```
Asn-Gln-Ile-Asp-Leu
 +   +   +   +   +
```

T-6

```
Asn-Ala-Ile-Lys
 +   +   +   -
```

T-7

```
Ser-Ala-Ile-Gly-Tyr-Thr-Pro-Trp-Ile-Gln-Cys-Asn-Val-Asp-Gln-Ser-
2081 2051 1541 885 588 666 349 239 944 421   +  291 683 162 203 201

Gly-Asn-Ser-Gln-Leu-Tyr-Gln-Val-Tyr-Ile-Cys-Val-Asp-Gly-Ser-Gly-
249 190 178 147 235 136  78 251 123 193   +  159  66  81  42  64

Ser-Ser-Leu-Ile-Glu-Cys-Pro-Ile-Phe-Pro-Gly
 34  43  22  15   5   +   4   4   4   4   2
```

T-8

```
Cys-Gly-Thr-Ser-Ile-Glu-Phe-Pro-Thr-Phe
 +  671 536 355 504 173 296 198 179 111
```

A

```
Pro-Asn-Ser-Pro-Tyr-Asp-Gln-Ser-Gln-Ile-Ser-Asp-Leu-Ile-Ser-Ser-
248 122 131 130  96  52  76  75  34  68  38  20  39  30  25  37

Met-Gln-Gln-Asn-Trp-Pro-Thr-Leu
 12   +   +   +   -   +   +   +
```

TH-1

```
Leu-Trp-Pro-Asn-Asn-Asn-Asp-Gly-Thr-Tyr-Pro-Ser-Asn-Cys-Asp-Pro-
989 450 582 391 437 440 167 306 287 263 306 148 159   +  101 199

Asn
119
```

TH-2

```
Phe-Trp-Ser-His-Glu-Trp-Glu-Lys-His-Gly-Thr
1446  +  766 128 604 274 864 521  59  59   +
```

C

```
Gln-Gln-Asn-Trp-Pro-Thr-Leu-Ala-Cys-Pro-Ser-Gly-Ser-Gly-Ser-Thr-
330 523 326   +   73 214 197 256   +  125  78 104  90  71  56  18

Phe-Trp-Ser-His-Glu-Trp-Glu-Lys-His-Gly-Thr
 36   +  32   9   +   -   +   +   +   +   +
```

G

```
Ser-Val-Leu-Thr-Asn-Gln-His-Ala-Tyr-Phe-Lys-Lys-Ala-Leu-Asp-Leu-
1207 1976 1643 1611 796 1163 209 519 1138 1318 1014 1085 1586 1108 418 1118

Lys-Asn-Gln-Ile-Asp-Leu-Leu-Ser-Ile-Leu-Gln-Gly-Ala-Asp-Ile-His-
784 313 687 888 330 484 957   +  266 558   +  521   +  219 489  88

Pro-Asp-Gly-Glu-Ser-Tyr-Asp-Leu-Val-Asn-Ile-Arg-Asn-Ala-Ile-Lys-
122  91 174  84  69  78  28 107  97  75  88  39  23  86 111  31

Ser-Ala-Ile-Gly-Tyr-Thr-Pro
 +   45  66  38  18  26  25
```

CH

```
Ile-Cys-Val-Asp-Gly-Ser-Gly-Ser-Ser-Leu-Ile-Glu-Cys-Pro-Ile-Phe-
437  +  363 109 286 216 207 156  88 105  94  38   +   48  41  39

Pro-Gly-Gly-Lys-Cys-Gly-Thr-Ser-Ile-Glu-Phe
 38  31  41   7   +  16  22  11   9   3   3
```

Fig. S1. *Sequencing data of intact RNase LE and fragments indispensable for sequence determination.* Numbers below the sequences refer to amount (pmol) of the corresponding phenylthiohydantoin derivatives. Peptides T-1, T-4, T-5 and T-6 were sequenced with the DABITC/PhSCN method. (+) unambiguous identification without a quantitative recovery; (−) not unambiguously identified